Which adolescents develop persistent substance dependence in adulthood? Using population-representative longitudinal data to inform universal risk assessment

M. H. Meier1*, W. Hall2,3, A. Caspi4,5,6, D. W. Belsky7,8, M. Cerdá9, H. L. Harrington4,5, R. Houts4,5, R. Poulton10 and T. E. Moffitt4,5,6

1 Department of Psychology, Arizona State University, Tempe, AZ, USA
2 University of Queensland Centre for Clinical Research, University of Queensland, Brisbane, Australia
3 National Addiction Centre, King’s College, London, UK
4 Department of Psychology and Neuroscience, Duke University, Durham, NC, USA
5 Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA
6 Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, King’s College London, UK
7 Department of Medicine, Duke University Medical Center, Durham, NC, USA
8 Social Science Research Institute, Duke University, Durham, NC, USA
9 Department of Emergency Medicine, Violence Prevention Research Program, University of California Davis, Davis, CA, USA
10 Dunedin Multidisciplinary Health and Development Research Unit, Department of Psychology, University of Otago, Dunedin, New Zealand

Background. To our knowledge, there are no universal screening tools for substance dependence that (1) were developed using a population-based sample, (2) estimate total risk briefly and inexpensively by incorporating a relatively small number of well-established risk factors, and (3) aggregate risk factors using a simple algorithm. We created a universal screening tool that incorporates these features to identify adolescents at risk for persistent substance dependence in adulthood.

Method. Participants were members of a representative cohort of 1037 individuals born in Dunedin, New Zealand in 1972–1973 and followed prospectively to age 38 years, with 95% retention. We assessed a small set of childhood and adolescent risk factors: family history of substance dependence, childhood psychopathology (conduct disorder, depression), early exposure to substances, frequent substance use in adolescence, sex, and childhood socioeconomic status. We defined the outcome (persistent substance dependence in adulthood) as dependence on one or more of alcohol, tobacco, cannabis, or hard drugs at ≥3 assessment ages: 21, 26, 32, and 38 years.

Results. A cumulative risk index, a simple sum of nine childhood and adolescent risk factors, predicted persistent substance dependence in adulthood with considerable accuracy (AUC = 0.80).

Conclusions. A cumulative risk score can accurately predict which adolescents in the general population will develop persistent substance dependence in adulthood.

Received 7 August 2015; Revised 14 October 2015; Accepted 26 October 2015

Key words: Longitudinal, risk assessment, substance dependence, universal screen.

Introduction

There is increasing interest in community-based, universal risk assessment to identify youth who either have a substance-use disorder or who will develop one in the future. Universal risk assessment, followed by appropriate intervention, could potentially reduce the population burden of disease associated with substance-use disorders. There are many risk-assessment tools that screen adolescents for current or future substance-use disorder (Clark et al. 2006; Vanyukov et al. 2009; Chung et al. 2012; Kirisci et al. 2013; Levy et al. 2014). None, to our knowledge, incorporate the key features of the most successful universal risk assessment tools to date, such as the Framingham risk score for cardiovascular disease (Wilson et al. 1998; D’Agostino et al. 2008). The Framingham risk score was developed using population-based samples, estimates total risk briefly and inexpensively by incorporating a relatively small number of well-established risk factors, and aggregates risk factors using a simple algorithm.
In the present study, we developed a risk score to identify adolescents in the general population who are at risk for persistent substance dependence. To maximize translation to practice in community settings, we incorporated the key features of the most successful universal risk assessment tools. That is, we used data from a population-representative longitudinal study; we selected a relatively small number of risk factors that have been shown in longitudinal studies to consistently and robustly predict substance dependence; and we aggregated these risks into a risk score using a simple algorithm, namely the sum of dichotomous risks. This summation approach draws on the large body of research showing that number of childhood risk factors predicts poorer mental and physical health in adulthood (i.e. cumulative risk) (Rutter, 1981; Sameroff et al. 1987; Felitti et al. 1998; Evans et al. 2013).

We evaluated the accuracy with which the cumulative risk score in adolescence predicted risk for substance dependence through young adulthood to early midlife. However, rather than predicting an adolescent’s risk for lifetime substance dependence, we predicted risk for severe, persistent substance dependence. Our rationale was that epidemiological studies show that the prevalence of lifetime substance dependence is quite high, and most people with substance dependence remit on their own without treatment (Heyman, 2013; Meier et al. 2013; Grant et al. 2015). Therefore, to avoid over-treating adolescents who would benefit from brief, harm-reduction interventions, and under-treating adolescents who might require more intensive intervention, we developed a population-based risk score that distinguishes those with the poorest long-term prognosis from those with a relatively good prognosis. In addition, rather than predicting risk for dependence on specific substances (e.g. alcohol vs. cannabis), we collapsed across substances in defining persistent forms of substance dependence. Our reasoning was that practitioners conducting universal screening (e.g. primary care physicians) want to assess risk for severe dependence on any substance, as opposed to risk for particular types of substance dependence. We further reasoned that the development of substance-specific risk assessment tools would result in the proliferation of risk assessments, thereby reducing implementation in practice.

Method

Participants

Participants are members of the Dunedin Multidisciplinary Health and Development Study, a longitudinal investigation of health and behavior in a complete birth cohort (Poulton et al. 2015). Study members (N = 1037, 91% of eligible births, 52% male) were all individuals born between April 1972 and March 1973 in Dunedin, New Zealand, who were eligible for the longitudinal study based on residence in the province at age 3 and who participated in the first follow-up assessment at age 3. The cohort represents the full range of socioeconomic status (SES) in the general population of New Zealand’s South Island and is primarily white (Moffitt et al. 2001). On adult health, the cohort matches the NZ National Health & Nutrition Survey (e.g. body mass index, smoking, general practitioner visits) (Poulton et al. 2006). Assessments occurred at birth and at ages 3, 5, 7, 9, 11, 13, 15, 18, 21, 26, 32, and, most recently, 38 years, when 95% of the living 1007 study members took part. At each assessment phase, study members were brought to the Dunedin Research Unit for a full day of interviews and examinations.

Ethics statement

The study protocol was approved by the institutional ethical review boards of the participating universities. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The Otago Ethics Committee approved each phase of the study. Informed consent was obtained from all study members.

Persistent substance dependence

Past-year substance dependence diagnoses were made using the Diagnostic Interview Schedule (DIS) following Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria (Robins et al. 1981, 1995; APA, 1987, 1994). We assessed alcohol, tobacco, and cannabis dependence at ages 21, 26, 32, and 38 and hard-drug dependence (e.g. heroin, cocaine) at ages 26, 32, and 38. DSM-III-R criteria were used at age 21 and DSM-IV criteria were used at ages 26, 32, and 38. We have previously compared prevalence rates of alcohol and cannabis dependence in the Dunedin Study with other representative studies of same-age respondents (Moffitt et al. 2010). The past-year prevalence of alcohol dependence was similar in the Dunedin Study, the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), and the National Comorbidity Survey (NCS; Moffitt et al. 2010). The past-year prevalence of cannabis dependence was slightly higher in the Dunedin Study than in representative U.S. surveys, but was similar to another longitudinal, population-representative survey of New Zealanders (Moffitt et al. 2010). The past-year prevalence of tobacco dependence in the Dunedin
Study (averaged across ages 18–38, 17%) was similar to NESARC participants aged 18+ (13%) (Grant et al. 2004). (NCS did not assess past-year tobacco dependence.) Direct comparison of rates of hard-drug dependence across studies is difficult due to differences between studies in the drugs included in this category. However, in general, rates of hard-drug dependence appear to be slightly higher in the Dunedin Study. To summarize, the prevalence of alcohol and tobacco dependence in the Dunedin Study is similar to other representative U.S. studies, but the prevalence of cannabis and hard-drug dependence is slightly higher in Dunedin. One potential explanation for this is that the prevalence of cannabis and hard-drug dependence is, indeed, higher in New Zealand. Another potential explanation for this pattern of findings is that Dunedin Study participants, interviewed repeatedly over the course of their lives, have learned to trust the study’s confidentiality guarantee, and are, therefore, more forthcoming about their illicit drug use.

We classified study members as persistently substance dependent if they were diagnosed with one or more of alcohol, tobacco, cannabis, or hard-drug dependence at 3+ assessment ages (ages 21, 26, 32, 38). For example, a study member could be considered persistently dependent if they were diagnosed with alcohol dependence at 3+ assessment ages (homotypic continuity). A study member could also be considered persistently dependent if they were diagnosed with tobacco dependence at age 21, followed by cannabis dependence at age 26 and hard-drug dependence at age 32 (heterotypic continuity). Of those classified as persistently dependent, 73% both diagnosed persistently for a single substance and across different substances. We chose a threshold of 3+ diagnoses to ensure that we were capturing individuals with severe, chronic dependence throughout adulthood.

We collapsed across substances in defining persistent dependence because practitioners want to predict risk for severe dependence on any substance, rather than dependence on a particular substance. This decision to collapse across substances is bolstered by evidence that (a) different substance-use disorders tend to co-occur (Table 1) (Krueger et al. 2002; Kendler et al. 2003; McGue et al. 2006), (b) a common liability underlies all substance-use disorders (Krueger et al. 2002; Kendler et al. 2003; McGue et al. 2006) and, (c) our results were similar across specific substances (Supplementary Tables S1–S4).

To be classified as persistently substance dependent, study members had to have been assessed for dependence at three of four assessment occasions. Ninety-three percent of the original 1037-member cohort were classified (910 had diagnostic data for four assessment occasions and 961 had data for three). Of those not classified, nearly half ($n = 37$) had either died or left the study before age 18 or had severe developmental disabilities that prevented their being interviewed with the DIS.

**Risk factors**

The nine childhood and adolescent risk factors are described in Table 2: SES, family history of substance dependence, conduct disorder, depression, early exposure to substances, frequent alcohol use, frequent tobacco use, frequent cannabis use, and male sex. We selected these particular risk factors because they (i) have been shown in longitudinal studies to consistently and robustly predict adult substance dependence, (ii) represent pre-specified domains of obvious interest (sociodemographic characteristics, mental health, and substance use), and (iii) have fairly natural cut-offs (Grant & Dawson, 1998; Chassin et al. 2004; Fergusson et al. 2007; Pardini et al. 2007; Brook et al. 2011; Hussong et al. 2011; Stone et al. 2012; Kendler et al. 2013).

**Statistical analysis**

We summed the nine dichotomous childhood and adolescent risk factors to produce a single cumulative risk index that allowed us to classify individuals as persistently substance dependent based on their number of risks. We evaluated predictive accuracy using the traditional performance measures: area-under-the-curve, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Choosing a higher threshold for number of risks results in lower sensitivity (more false negatives) but higher specificity (fewer false positives). We plotted sensitivity against 1 – specificity for every value of the index, yielding a receiver-operating characteristic (ROC) curve. The area under the curve (AUC) provides a measure of predictive accuracy that reflects the probability of correctly classifying a randomly selected pair of individuals in which one has persistent substance dependence and the other does not. The AUC can take on any value between 0.50 (indicating chance prediction) and 1.00 (indicating perfect prediction). AUC values of 0.54, 0.64, and 0.71 correspond to Cohen’s $d$ values of 0.20, 0.50, and 0.80, and reflect small, medium, and large effects, respectively (Rice & Harris, 2005).

**Results**

The prevalence of persistent adult substance dependence to age 38 in this population-representative cohort was 19% ($n = 183$). Table 3 shows that each childhood and adolescent risk factor significantly predicted persistent substance dependence in adulthood, and frequent tobacco use in adolescence was the best predictor (Table 3, left panel; AUC = 0.74). Adolescent tobacco use remained a top predictor even when tobacco
Table 1. Prevalence of persistent (3+ diagnoses between ages 21 and 38) and lifetime (at least one diagnosis between ages 21 and 38) substance dependence by study member category

<table>
<thead>
<tr>
<th>Substance dependence</th>
<th>All adults (N = 961)</th>
<th>Persistent alcohol dependence (N = 40)</th>
<th>Persistent tobacco dependence (N = 96)</th>
<th>Persistent cannabis dependence (N = 27)</th>
<th>Persistent hard-drug dependence (N = 10)</th>
<th>Persistent substance dependence (N = 183)</th>
<th>Persistent dependence on substances excluding tobaccoa (N = 83)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistent alcohol dependence</td>
<td>4.17</td>
<td>100.00</td>
<td>19.79</td>
<td>25.93</td>
<td>30.00</td>
<td>21.86</td>
<td>48.19</td>
</tr>
<tr>
<td>Persistent tobacco dependence</td>
<td>10.02</td>
<td>47.50</td>
<td>100.00</td>
<td>48.15</td>
<td>50.00</td>
<td>53.04</td>
<td>38.55</td>
</tr>
<tr>
<td>Persistent cannabis dependence</td>
<td>2.81</td>
<td>17.50</td>
<td>13.54</td>
<td>100.00</td>
<td>30.00</td>
<td>14.75</td>
<td>32.53</td>
</tr>
<tr>
<td>Persistent hard-drug dependence</td>
<td>1.07</td>
<td>7.69</td>
<td>5.32</td>
<td>11.54</td>
<td>100.00</td>
<td>5.68</td>
<td>12.35</td>
</tr>
<tr>
<td>Lifetime alcohol dependence</td>
<td>31.84</td>
<td>100.00</td>
<td>65.62</td>
<td>70.37</td>
<td>60.00</td>
<td>73.22</td>
<td>85.54</td>
</tr>
<tr>
<td>Lifetime tobacco dependence</td>
<td>34.03</td>
<td>72.5</td>
<td>100.00</td>
<td>85.19</td>
<td>70.00</td>
<td>89.07</td>
<td>75.90</td>
</tr>
<tr>
<td>Lifetime cannabis dependence</td>
<td>16.34</td>
<td>55.0</td>
<td>42.71</td>
<td>100.00</td>
<td>90.00</td>
<td>53.55</td>
<td>74.70</td>
</tr>
<tr>
<td>Lifetime hard-drug dependence</td>
<td>6.56</td>
<td>35.00</td>
<td>21.87</td>
<td>48.15</td>
<td>100.00</td>
<td>26.78</td>
<td>45.78</td>
</tr>
</tbody>
</table>

This table shows the prevalence of the disorders listed in the rows given the group listed in the columns. For example, of all adults in the cohort (column 1, N = 961), 4.17% had persistent alcohol dependence (row 1). As another example, of those with persistent alcohol dependence (column 2, N = 40), 47.50% had persistent tobacco dependence (row 2).

* Even after excluding tobacco dependence from the criteria for persistent substance dependence, 38.55% of the group with persistent substance dependence also met criteria for persistent tobacco dependence (last column, second row). This is because many of those who were persistently dependent on alcohol, cannabis, or hard drugs were also persistently dependent on tobacco, and 75.90% (last column, sixth row) had been dependent on tobacco at some point in their life between ages 21 and 38.
<table>
<thead>
<tr>
<th>Risk</th>
<th>Respondent</th>
<th>Description</th>
<th>Study member's age(s) at assessment</th>
<th>Brief screen adaptation of risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Low family SES</td>
<td>Parents</td>
<td>The highest of father’s or mother’s occupation using a 6-point scale for New Zealand (Elley &amp; Irving, 1976). Repeated measures were averaged (Wright et al. 1999). Study members were divided into two groups: high or intermediate SES (manager or physician, secretary or electrician) and low SES (cashier or textile machine operator)</td>
<td>Birth–15</td>
<td>Not included in the brief screen, because this measure is based on data aggregated across 15 years</td>
</tr>
<tr>
<td>(2) Family history of substance dependence</td>
<td>Study member and parents</td>
<td>The Family History Screen (Weissman et al. 2000) was used to obtain the proportion of family members, across three generations, with a diagnosis of substance dependence (alcohol or drug dependence) (Milne et al. 2009a, b). Study members with a proportion of 30% or more were classified as having a family history of substance dependence</td>
<td>32</td>
<td>Maternal reports of alcohol problems for six family members: study members’ biological parents and all four grandparents. Mothers answered the following questions: ‘Has ___ ever had any treatment or been in hospital for drinking?’ ‘Has ___ ever had alcoholism?’ ‘Has ___ ever had a drinking problem or did other people think he/she had a drinking problem?’ (Odgers et al. 2007). An affirmative answer to any question was deemed positive for family history</td>
</tr>
<tr>
<td>(3) Childhood conduct disorder</td>
<td>Study member</td>
<td>Past-year conduct disorder was assessed using the Diagnostic Interview Schedule for Children (Costello et al. 1982) at ages 11, 13, and 15 and the Diagnostic Interview Schedule (Robins et al. 1981) at age 18. Diagnoses were based on DSM-III criteria at the younger ages and DSM-III-R criteria at age 18. Conduct disorder criteria at each assessment phase were scored to be consistent with DSM-IV (Moffitt et al. 2001)</td>
<td>11, 13, 15, and 18</td>
<td>One symptom was taken from the age 18 conduct disorder assessment. We selected the symptom that correlated most highly with the full diagnosis: breaking in</td>
</tr>
<tr>
<td>(4) Childhood depression</td>
<td>Study member</td>
<td>Past-year depression was assessed using the Diagnostic Interview Schedule for Children (Costello et al. 1982) at ages 11, 13, and 15 and the Diagnostic Interview Schedule (Robins et al. 1981) at age 18. Diagnoses were based on DSM-III criteria at the younger ages and DSM-III-R criteria at age 18</td>
<td>11, 13, 15, and 18</td>
<td>One symptom was taken from the age 18 depression assessment. We selected the symptom that correlated most highly with the full diagnosis: fatigue or loss of energy</td>
</tr>
<tr>
<td>(5) Early exposure to substances</td>
<td>Study member</td>
<td>Use of drugs (e.g. inhalants, cannabis) or use or purchase of alcohol on multiple occasions over the past year at age 13, age 15, or both (Odgers et al. 2008)</td>
<td>13 and 15</td>
<td>Same</td>
</tr>
<tr>
<td>(6) Adolescent frequent alcohol use</td>
<td>Study member</td>
<td>Study members reported on their frequency of alcohol use over the past year at age 18. Study members who reported using alcohol on 5+ days per week were considered frequent alcohol users</td>
<td>18</td>
<td>Same</td>
</tr>
</tbody>
</table>
dependence was excluded from the outcome (Table 3, right panel; AUC = 0.69), which was unsurprising given the high rates of co-morbidity between tobacco dependence and dependence on other substances (Table 1).

The cumulative risk index (mean = 1.78, s.d. = 1.50) predicted persistent substance dependence in adulthood with considerable accuracy. The ROC analysis revealed an AUC of 0.80, a large effect, meaning that we had an 80% probability of correctly predicting, from a randomly selected pair of adolescents, which adolescent would have persistent substance dependence in adulthood. Results were similar when tobacco dependence was excluded from the outcome (Table 3, right panel; AUC = 0.81) and when predicting persistent dependence on each substance individually (Supplementary Tables S1–S4).

The prevalence of persistent substance dependence increased markedly as a function of number of childhood and adolescent risks (Fig. 1a, striped bars): 3% of adolescents with 0 risk factors, 27% of adolescents with 3 risk factors, and 74% of adolescents with 6+ risk factors had persistent substance dependence as adults.

Table 4 shows sensitivity, specificity, PPV, NPV, and overall classification accuracy of the cumulative risk index as a function of number of risks. Overall accuracy was greatest at a cut-off on the cumulative risk index of 4+ to 5+ risks.

**Moderation by sex**

Male sex was included as a risk factor in the cumulative risk index, but to test the possibility that the cumulative risk index is a more accurate predictor of persistent substance dependence for one sex, we examined sex as a moderator. First, we tested whether there were sex differences in the associations between each risk factor in the cumulative risk index and persistent substance dependence. There was no evidence that sex moderated any of these associations. Next, we removed male sex as a risk factor in the cumulative risk index and recomputed the AUC for the cumulative risk index separately by sex. The cumulative risk index (without sex as a risk factor) predicted persistent substance dependence similarly well for girls (AUC = 0.81) and boys (AUC = 0.78). Moreover, the cut-off score that maximized overall classification accuracy was 4+ to 5+ risks for girls and 3+ risks for boys, v. 4+ to 5+ risks for both girls and boys when male sex was included as a risk factor in the cumulative risk index. Findings suggest that the cumulative risk index predicts persistent substance dependence similarly for girls and boys. A practical advantage of including male sex as a risk factor in the cumulative risk index is that it equates the cut-off score for girls and boys.
Sensitivity analyses
We tested whether prediction could be improved by adding another risk factor to the cumulative risk index – either low childhood self-control or childhood maltreatment. These risk factors were selected because they have been shown to predict risk for substance dependence and because they were available in the data-set. Both low childhood self-control and childhood maltreatment were associated with increased risk of persistent substance dependence, and they predicted persistent substance dependence with similar accuracy (Supplementary Table S5; AUCs = 0.55 and 0.57, respectively). Adding these risk factors to the cumulative risk index did not improve accuracy. When either risk factor was added, the cumulative risk index predicted persistent substance dependence with an AUC of 0.79.

Next we tested the effects of dropping a risk factor from the cumulative risk index. Table 5 shows that the AUC did not drop substantially with the exclusion of any single risk factor, with exception of frequent tobacco use in adolescence.

Finally, we tested the robustness of the cumulative risk index by reducing the threshold for ‘persistent’ substance dependence from 3+ to 2+ diagnoses across assessment ages. The cumulative risk index was about as accurate when predicting 2+ substance dependence diagnoses from ages 21 to 38 (AUC = 0.77) as when predicting 3+ diagnoses (AUC = 0.80). The cumulative risk index was also fairly accurate when predicting 1+ diagnoses from ages 21 to 38 (AUC = 0.76).

The cut-off score on the cumulative risk index that maximized overall classification accuracy dropped from 4–5+ risks to 3–4+ risks to 2+ risks for predicting 3+, 2+, and 1+ dependence diagnoses, respectively, from ages 21 to 38. This is not surprising given the dose-response association between number of childhood and adolescent risk factors and persistence of substance dependence. For example, the mean number of substance dependence diagnoses from ages 21 to 38 increased in a fairly linear fashion as a function of number of childhood and adolescent risk factors (Fig. 1b). Fig. 1a shows that cohort members with a greater number of risk factors were at higher risk of dependence, regardless of how we defined the dependence outcome (1+, 2+, or 3+ diagnoses). The vast majority of individuals with 0 risks never developed dependence (79%), whereas this was true of only 6% of those with 4+ risks.

Comparison with other risk screens
Both NIDA’s Quick Screen and NIAAA’s Alcohol Screening Guide for Children and Adolescents rely exclusively (NIDA) or heavily (NIAAA) on assessing frequency of drug use and/or alcohol use. Table 3 and Supplementary Tables S1–S4 show that by expanding risk assessment beyond frequency of drug and alcohol use, prediction was improved. For example, frequent cannabis use in adolescence predicted persistent cannabis dependence in adulthood with an AUC of 0.83 (Supplementary

### Table 3. Relative risk of persistent substance dependence in adulthood given each childhood and adolescent risk factor

<table>
<thead>
<tr>
<th>Childhood and adolescent risks</th>
<th>Persistent dependence on alcohol, tobacco, cannabis, and hard-drugs (19%)</th>
<th>Persistent dependence on substances except tobacco (9%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR</td>
<td>95% CI</td>
<td>p</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Low family socioeconomic status</td>
<td>1.73</td>
<td>1.32–2.27</td>
</tr>
<tr>
<td>Family history of substance dependence</td>
<td>2.62</td>
<td>2.03–3.37</td>
</tr>
<tr>
<td>Childhood conduct disorder</td>
<td>3.20</td>
<td>2.50–4.09</td>
</tr>
<tr>
<td>Childhood depression</td>
<td>2.05</td>
<td>1.58–2.67</td>
</tr>
<tr>
<td>Early exposure to substances</td>
<td>2.91</td>
<td>2.24–3.79</td>
</tr>
<tr>
<td>Adolescent frequent alcohol use</td>
<td>2.30</td>
<td>1.49–3.54</td>
</tr>
<tr>
<td>Adolescent frequent tobacco use</td>
<td>5.41</td>
<td>4.00–7.31</td>
</tr>
<tr>
<td>Adolescent frequent cannabis use</td>
<td>4.25</td>
<td>3.22–5.61</td>
</tr>
<tr>
<td>Male</td>
<td>1.54</td>
<td>1.18–2.02</td>
</tr>
<tr>
<td>Cumulative risk index</td>
<td>0.80</td>
<td></td>
</tr>
</tbody>
</table>

RR, Relative risk; CI, confidence interval; AUC, area under the curve.

The risk factors in this table were modestly correlated (r’s ranged in absolute magnitude from 0.00 to 0.33). Note that AUC is a better indicator of classification accuracy than relative risk (Pepe et al. 2004).
The cumulative risk index was more accurate in predicting persistent dependence on each substance than frequency of alcohol or drug use alone, with one exception. Frequent tobacco use in adolescence predicted persistent tobacco dependence as well as the cumulative risk index (Supplementary Table S2).

**Brief screen adaptation**

We adapted our risk measures for a brief screen to see whether the cumulative risk index could be used in community settings by practitioners with limited time to conduct the detailed risk assessments in our research study. Table 2 shows how we adapted each measure, and Supplementary Table S6 shows the brief screen. The brief screen (i.e., a cumulative risk index based on the sum of the eight adapted risks) performed nearly as well the cumulative risk index based on our more detailed risk assessments (AUC = 0.79 vs. 0.80, respectively). Like the cumulative risk index based on our more detailed risk assessments, the brief screen was more accurate in predicting persistent...
dependence on each substance than frequency of alcohol or drug use alone with one exception. Frequent tobacco use in adolescence predicted persistent tobacco dependence as well as the brief screen. We report additional information on the accuracy of this brief-screen adapted risk index in Supplementary Table S7 and Supplementary Fig. S1.

Discussion

This report advances knowledge by suggesting answers to three recently posed questions about screening adolescents for risk of substance-use disorders (Subramaniam & Volkow, 2014). The first question is: How can we combine multiple risk factors to estimate an adolescent’s risk? Current screening tools rely heavily on assessing adolescent substance use, yet other risk factors, such as psychiatric disorder, also predict risk (Subramaniam & Volkow, 2014). In this report, we showed that summing a small set of dichotomous risks into a single cumulative risk index is a clinically useful way to integrate risk factors and accurately predict persistent substance dependence. Many studies have shown that cumulative childhood risk is associated with adult mental and physical health problems (Rutter, 1981; Sameroff et al. 1987; Felitti et al. 1998; Evans et al. 2013). A recent study even showed that cumulative risk distinguishes those with persistent alcohol problems from those with time-limited alcohol problems (Copeland et al. 2012). The current report extends this work to suggest how cumulative risk could be used as an actuarial risk assessment tool in community settings to accurately predict persistent substance dependence in the general population.

The second question we address is: Who is at risk? This report provides initial population-representative estimates of risk that can be used to gauge an individual adolescent’s likelihood of having persistent substance dependence in adulthood. For example, 3% of adolescents with 0 risks, 27% of adolescents with 3 risks, and 74% of adolescents with 6+ risks developed persistent substance dependence in adulthood. In the future, practitioners may use risk estimates such as these (aggregated across more population-representative studies like ours) to make actuarial judgments and referrals to treatment. One may even envisage members of the general population calculating their risk for persistent substance dependence on their own, just as they now can calculate their risk for heart attack (http://cvdrisk.nhlbi.nih.gov/calculator.asp).

<table>
<thead>
<tr>
<th>Risk factor removed from the cumulative risk index</th>
<th>AUC for the cumulative risk index</th>
</tr>
</thead>
<tbody>
<tr>
<td>No risk factors removed</td>
<td>0.80</td>
</tr>
<tr>
<td>Low family socioeconomic status</td>
<td>0.80</td>
</tr>
<tr>
<td>Family history of substance dependence</td>
<td>0.79</td>
</tr>
<tr>
<td>Childhood conduct disorder</td>
<td>0.79</td>
</tr>
<tr>
<td>Childhood depression</td>
<td>0.79</td>
</tr>
<tr>
<td>Early exposure to substances</td>
<td>0.79</td>
</tr>
<tr>
<td>Adolescent frequent alcohol use</td>
<td>0.80</td>
</tr>
<tr>
<td>Adolescent frequent tobacco use</td>
<td>0.75</td>
</tr>
<tr>
<td>Adolescent frequent cannabis use</td>
<td>0.80</td>
</tr>
<tr>
<td>Male</td>
<td>0.78</td>
</tr>
</tbody>
</table>

AUC, Area under the curve.
The third question we address is: How should we decide the appropriate level of intervention for an at-risk adolescent? By triaging adolescents into different levels of intervention based on their risk, we might avoid under- or over-treatment and reduce the costs of treatment. The cumulative risk index lends itself to triaging, because we documented a dose-response association between number of childhood and adolescent risk factors and persistence of substance dependence. That is, individuals with more risks had more persistent substance dependence. This dose-response association suggests that adolescents with more risks may require more intensive intervention, whereas individuals with fewer risks may benefit from brief interventions. Although more work will be needed to determine the cut-off scores on the cumulative risk index that decide level of care, we illustrate one possible way that the cumulative risk score could be used.

Adolescents with 4+ risks might be candidates for intensive intervention. A cut-off score of 4+ risks maximized overall accuracy in predicting persistent dependence in adulthood. Moreover, specificity at 4+ risks was high (93%), ensuring that costly interventions go to the small portion of the population (13.6%) that really need it (Table 4). Overtreatment of adolescents with 4+ risks is unlikely, as 94% were diagnosed with dependence at least once between ages 21 and 38 and 82% were diagnosed at least twice. These adolescents may benefit from broad-based interventions (i.e. interventions that target a variety of risk factors, not just substance use) because of the number and variety of their accumulated risks. Interventions need not focus exclusively on the small set of predictors studied here. These predictors were selected for the explicit purpose of efficiently identifying those at highest risk. Intervention design should be guided by the literature on effective treatments that target mutable risks for a variety of problem behaviors.

Brief interventions might be offered to adolescents with 2 and 3 risk factors. These adolescents were also at increased risk for persistent substance dependence, but compared with adolescents with 4+ risks, their struggles with dependence were more likely to be time-limited (Supplementary Fig. S2). The screening itself could serve as the basis for a brief intervention, as the cumulative risk index provides a powerful means of communicating to adolescents their overall risk in clear and understandable terms. Adolescents with 2 risks were seven times more likely than their peers with 0 risks to struggle with persistent dependence through early midlife. Adolescents with 3 risks were nine times more likely than their peers with 0 risks to have persistent dependence. For some adolescents, information about their risk status might motivate behavior change towards a lower risk category and prevent initiation or escalation of substance use. For adolescents who have 2 or 3 risk factors, and at least one of those risk factors includes frequent substance use, feedback about risk status might be combined with brief motivational interviewing and/or harm-reduction strategies focused on substance use. For adolescents with symptoms of depression or conduct problems, brief interventions targeting these symptoms might be appropriate (e.g. behavioral activation for depression or brief behavioral parenting strategies for disruptive behavior).

Our results may well prove useful in designing and evaluating intensive prevention/early intervention programs in the future. Researchers planning these programs need to know approximately what proportion of their treatment group would develop persistent dependence. This information could be used to calculate the ‘number needed to treat’ to prevent one case – a measure that is increasingly used to gauge the effectiveness of an intervention. For example, if adolescents with 4+ risks were the target of intervention, 60% would otherwise develop persistent substance dependence in adulthood (Table 4, PPV). Given a perfectly effective intervention, the ‘number needed to treat’ to prevent one case would be 1.67 \((1/0.60 = 1.67)\) (Cook & Sackett, 1995). As the proportion of the treatment sample who would otherwise develop persistent dependence decreases, the number needed to treat increases. Thus, prevention/early intervention studies that cast too wide a net in defining their treatment group have, from the start, limited their treatment’s effectiveness.

The results of this study should be interpreted in the context of its limitations. First, although we selected some of the best predictors of substance dependence, prediction might be marginally improved by adding childhood and adolescent risks. Our sensitivity analyses showed that, with exception of adolescent frequent tobacco use, adding or subtracting a predictor did not make a difference in accuracy of the cumulative index. Moreover, each additional predictor yields diminishing returns while lengthening assessment (Ware, 2006). Substituting different predictors from the ones we examined here could lead to improvements in prediction, but multiple datasets will be needed to fully explore this possibility to avoid overfitting the model to this cohort. Substantial improvement in prediction may be unlikely however, because prediction is already quite good. A more promising strategy for improving upon prediction might be to include more predictors in models that could reveal combinations of risk factors that predict an especially high risk of persistent substance dependence (e.g. decision tree, cluster analysis, or neural network models).
Predicting which adolescents develop persistent adult substance dependence

A second limitation is that our findings are based on a single New Zealand cohort and require replication in independent samples. We view this report as proof of concept for using a cumulative risk index as an actuarial tool. The next steps include testing and refining the cumulative risk index in contemporary cohorts and in special populations, such as Native American adolescents, as well as testing the accuracy of the cumulative risk index in predicting persistent substance-use disorder, as defined by DSM-5.

In summary, we developed a universal screen for persistent substance dependence that (1) is based on population-representative data, (2) estimates total risk briefly and inexpensively by incorporating a relatively small number of well-established risk factors, and (3) aggregates risk factors using a simple algorithm. Although findings are preliminary, they suggest that we can predict with considerable accuracy which adolescents in the general population will struggle with persistent substance dependence in adulthood. The cumulative risk index may be simpler to use, less costly, and more accurate (AUC = 0.80) than more comprehensive screens, and these features are important considerations in universal screening. For example, an extensive profile of neuropsychosocial risk, including measures of brain and cognitive function, predicted adolescent binge drinking with an AUC of 0.75 (Whelan et al. 2014). We also showed that the cumulative risk index compared favorably to current risk assessment approaches, which narrow in on substance use as the primary risk indicator. Moreover, an adapted version of the cumulative risk model for use in community settings (a version that simply summed readily obtained adolescent risks) yielded an AUC of 0.79. Additional research is needed to validate the cumulative risk index, evaluate its practical utility, and address potential ethical issues that may be raised by screening adolescents for persistent substance dependence (Carter & Hall, 2011; Hall et al. 2015). The results presented here represent a first step toward establishing population-representative estimates of risk for persistent substance dependence that may be useful in research and practice.

Supplementary material

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291715002482.

Acknowledgements

We thank the Dunedin Study members, their families, Unit research staff, and Study founder Phil Silva. The Dunedin Multidisciplinary Health and Development Research Unit is supported by the New Zealand Health Research Council and New Zealand Ministry of Business, Innovation, and Employment (MBIE). This research received support from US-National Institute on Aging (AG032282, AG048895); UK Medical Research Council (MR/K00381X); and ESRC (ES/M010309/1). Additional support was provided by the Jacobs Foundation. Wayne Hall was funded by a National Health and Medical Research Council Australia Fellowship. Daniel W. Belsky was supported by the National Institute on Aging (T32-AG000029).

Declaration of Interest

None.

References


